DOCUMENT RESUME

ED 098 456 95 CG 009 198

AUTHOR Medin, Douglas L.
TITLE Learning to Learn in Children.
INSTITUTION Rockefeller Univ., New York, N.Y.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE [72]
GRANT OEG-2-71-0532
NOTE 20p.

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS *Child Development; Children; Cognitive Processes; *Discrimination Learning; *Learning; Research Projects; *Transfer of Training

ABSTRACT

Seven main experiments and three pilot studies were conducted to develop and test theories of transfer of training in children. Initial tests failed to support an incentive model for learning and transfer, but a new model given strong emphasis to the role of context in learning was developed which accounted for a wide range of learning and transfer data. Although this was clearly the most salient result, other work revealed that: (1) learning rules when there are exceptions to rules may be much more difficult than one might have been led to imagine; (2) simply solving any problem (learning the rule) may underlie the basis of the easy-to-hard transfer effects in children; (3) tests to measure dimensional preferences may create them; (4) young children may use more than a single dimension in making choices when decisions can be considered sequentially; (5) dimensional processing skills may increase with age in ways that are more significant than simply changes in the allocation of attention; and (6) some claims concerning changes in mediating skills with age were most likely premature. (Author)

CG



# Final Report on Office of Education Grant (OEG2 71 0532)

# "Learning to Learn in Children"

Principal Investigator: Douglas L. Medin

The Rockefeller University

York Avenue & 66th Street

New York, New York 10021

## I. Summary of Progress:

The overall purpose of the project research was to gather information relevant to transfer phenomena such as learning to learn effects in children. The long term goal toward which this project represents only an initial effort is to develop a theoretical account of the transition between slow and rapid learning rates (learning to learn).

The initial specific aims consisted of two major thrusts. One was to provide tests of the applicability of an incentive model for learning to learn (Medin, 1972). The major novelty in the incentive model was that attention was not assumed to be a fixed quantity but modifiable both in direction and amount by experience. A second closely related but broader goal was to attempt to examine possible sources for increases in learning rates such as (1) learning not to make premature choices, (2) learning to anticipate rewards, (3) learning to selectively attend to relevant task dimensions, (4) acquiring complex hypotheses or organizing strategies to mediate performance.

Several factors quickly led to a shift in focus in the research. First, some initial tests of the incentive model failed to provide any support for it and at the same time it became apparent that attempts to conduct long-term experiments in public schools are much more difficult than originally envisioned.

A slight modification in emphasis was produced by the research on attention which led to a series of experiments on dimensional preferences and the interpretation of developmental changes in learning rates. These experiments showed that (1)

situations designed to measure preferences could serve to create preferences, (2) children who are unlikely to use more than one dimension in making a decision can and do use multi-dimensional information when decisions can be made sequentially, (3) older children differ from younger ones not simply in how they allocate their attention but also in that older children become more sophisticated in perceptually processing information along all but the simplest dimensions.

However, by far the primary motivation for shifting specific aims while keeping the overall goal intact were some promising theoretical developments with respect to understanding within-problem learning in children. To be more specific, most theories of children's discrimination learning are applicable to a single experimental paradigm (simultaneous discrimination) and make either no predictions or incorrect predictions for other paradigms (a notable exception is Spiker's 1963 stimulus interaction hypothesis). A model taking into account contextual factors in learning was developed by the principal investigator which accounts for performance in a very broad range of test paradigms. In many situations where investigators have assumed that their data implied that different styles of learning must be operating in different groups of subjects, the context theory handles the results with a single learning process. The main point is that this theory may be more promising for interpreting transfer and learning to learn phenomena than more direct attacks on the problem have been.

During the project period eight major experiments and several

pilot studies were completed. Most of the data have been analyzed and will be reported in one form or another. At this time two papers have been published, two presented at conferences or professional meetings, one paper is in press, another has been submitted for publication, and at least two other papers are in preparation.

Although the theoretical efforts and empirical research took somewhat different courses than anticipated, overall, considerable progress has been made in achieving the goals outlined. Memory factors such as the effects of context on retrieval of information appear as a major factor in children's learning. In addition some earlier concepts of attention in children can be rejected, and some new approaches to studying dimensional processing were developed. Also there emerged some sharp empirical and theoretical constraints concerning when "conceptual learning" may be observed or inferred. Particular details will be discussed in the following Results section.

II. Principal Results:

A) Incentives and Discriminative Performence.

1. Effects of noncontingent success or failure on same-different judgments. This experiment studied the effects of failure and success on a sameness-difference task. The prediction of the incentive model was that children with a low expectation of success might perform more poorly than children with a higher expectation and that these differences would be attributable primarily to children making premature judgments (Vurpillot, 1968).

In the initial experiment 32 kindergarten children were randomly assigned to one of two groups. Group A received: (1) noncontingent failure (20% chance of being correct), (2) 24 same-difference trials, (3) noncontingent success (80% chance of being correct), and (4) an additional 24 same-difference trials. Group B had the same treatment except that the order of the success and failure experiences was reversed. The noncontingent task consisted of a situation where the experimenter held up two cards with pictures on them and asked "which one am I thinking of?" Different pictures were used for each of the 20 trials given in a particular phase of the experiment. The experimenter said yes or no on a predetermined number of trials regardless of the child's guess. The same-different materials were pairs of dogs, dolls, or moon faces and were either identical in shape or differed in one or two features (e.g., collar vs. no collar in the case of dogs).

It was anticipated that the first noncontingent task would establish an expected level of success and that the second group of noncontingent trials would provide the "success" or "failure" experience. Consequently, it was anticipated that Group A would perform better than Group B on the second sameness-difference task and that this advantage would show up on the "different" rather than the "same" trials.

The results did, in fact, show a significant advantage for Group A over Group B on the "different" trials of the same-difference task. However, Group A also showed an advantage (though not reliable) on the first same-difference trials which

raised the possibility that the significant effect might be due to a sampling error.

A second experiment was conducted with a new group of subjects as a partial replication of the first study. The paradigm differed in that two consecutive noncontingent tasks were given prior to the same-difference trials. The experiences on the two tasks were given either in the order of high to low reward probability (failure) or low to high reward probability (success). Twenty-four kindergarten subjects were assigned randomly to either the failure or success group.

The results showed that subjects averaged approximately 30% errors on the same-different task but no group differences appeared. This suggests that the effect in the first experiment was a sampling error or at least the the magnitude of the effect of failure or success experience is not very large, especially with respect to expectations arising from the incentive model.

2. Reward and transfer after discrimination learning. This experiment was designed to provide a more direct test of the incentive model for learning to learn. Thirty kindergarten children were given training to a criterion on two concurrently presented color discriminations in a child version of the Wisconsin General Test Apparatus. Following this, they were shifted to a condition where the two formerly incorrect stimuli were paired and the two formerly correct stimuli were paired to create two new concurrently presented discriminations.

The incentive model predicted that transfer learning would occur more rapidly when the two correct objects were paired than when the previously incorrect objects were given. Contrary

to the theory, there were no differences during transfer and subjects learned quite rapidly (2.54 mean errors vs. 2.62 mean errors). Once a child learned the "rules of the game" it did not matter whether particular objects had been correct or incorrect previously. This empirical conclusion led to the next line of work.

B. Within Problem Learning and Shift Performance

1. Easy-to-hard transfer. If initial learning is the major obstacle to transfer across problems (as in learning set paradigms), then training on an easy discrimination may facilitate transfer to a hard discrimination in comparison to training solely on the hard discrimination. This easy-to-hard transfer effect has already been shown in children (Spiker, 1959).

This experiment was designed to replicate and partially extend Spiker's findings. Subjects were 4-year-olds (mean age 50 months) obtained from nurseries of a local Foundling Home. All subjects received 4 days of training for 30 trials each day on discrimination problems in a WGTA. The hard problem involved a discrimination between two purple squares differing by one Munsell brightness step (Values 4 & 5). The easy problem was either a form discrimination or a discrimination between two purple squares differing by three Munsell brightness steps (Values 3 & 6). Eight subjects were randomly assigned to one of four groups: (1) Hard. Training was given for all four days on the difficult brightness discrimination. (2) Easy-to-Hard. The initial two days of training were given on the easy brightness discrimination followed by two days training on the hard brightness discrimination. The reinforcement combinations were con-

sistent so that if Value 3 had been correct for the easy problem, then Value 4 would be correct for the hard problem. (3) Easy-to-Hard Reversal. This group was trained and transferred like the Easy-to-Hard group except that during transfer the reward relationships were reversed (i.e., if Value 3 were correct earlier then Value 5 rather than Value 4 would be correct). (4) Easy Irrelevant. This group was trained on an easy form discrimination for two days and then switched to the hard brightness discrimination.

The results showed that the Hard group's performance was inferior to that of the other three groups on the last day of testing. The Easy-to-Hard group performed best followed by the Easy Irrelevant group. Not enough subjects were run in this pilot experiment to distinguish between groups other than to note the poor performance of the Hard group. The main effect was that solving an easy problem of any kind led to solving the difficult discrimination. The transfer effects were essentially in terms of whether or not a subject learned at all rather than in level of performance. Only two of the eight subjects in the Hard group showed above chance responding on the difficult discrimination after four days of training, while all subjects solving an easy discrimination (except two subjects in the Easy-to-Hard Reversal group) showed above chance performance on the final day of the difficult discrimination. This tends to support the conclusion of Reese (1963) that solving a single discrimination problem is often sufficient to produce effective learning set performance in children.

2. Learning rules when there are exceptions to rules. Although the concepts usually studied in laboratories are "perfect" rules, in many ordinary experiences rules must be formed which involve a few exceptions. This experiment examined learning rules with exceptions in kindergarten and third grade children.

All subjects were given two consecutive verbal discrimination learning tasks. For each task the stimuli were pictures which fell into one of two categories (i.e., animals or clothing, or toys or foods). The categories consisted of 12 instances and were paired to produce 12-item lists. Three groups were run at each age level: (1) The Rule group was trained on the first task where a single rule (all animals are correct) could mediate perfect transfer. (2) The Exception group was trained where a rule would work but there were 2 exceptions (i.e., animals except fox and lion are correct). (3) A Random group where six randomly selected animals and six items of clothing were correct. Initial training was given for 96 trials or until a learning criterion was met. All three groups were then trained on a second task with two new categories forming the lists and for which there was a rule having two exceptions. Twelve subjects were randomly assigned to each of the three groups at each of the two grade levels for a total of 72 subjects.

Overall, the 3rd graders learned faster than the kindergarten subjects in all conditions. Secondly, on the first task the Rule groups made significantly fewer errors (about half as many) as the other groups. But most strikingly there was vir-

tually no evidence of any rule learning in any cases where there were exceptions to the rule. This was true for all 3 conditions at both age levels. First list learning produced no differential effects on list two. The Rule groups made 78% errors on the exceptions on the first run through the second list (compared to a 50% chance level) suggesting that subjects were trying to form rules but apparently as few as 2 exceptions in 12 instances were enough to abolish all r 'e learning. These results need to be followed up, of course, but they imply that rule formation with exceptions may be distinctly different from learning all-or-none rules.

3. Subproblem analysis of discrimination shift learning. Since this paper has already been published (Medin, 1973a), only a brief summary of it will be given here. A commonly used paradigm for assessing developmental changes in mediation involves comparisons of reversal and nonreversal shifts. In particular Tighe, Glick and Cole (1971) suggested on the basis of a subproblem analysis of nonreversal shift performance that 4-year-olds solved problems in a paired-associate manner while 10-year-olds treated subproblems as a single problem.

The paper first demonstrates that this claim is not justified and that age changes in performance may have little to do with changes in mediation. A general analysis of performance on subproblems appropriate to a broad range of attention models was developed in the paper. This analysis leads to the rejection of current attention models for the behavior of rats, pigeons, turtles, and monkeys but not for young children. The

major unanswered question concerns what happens when it is assumed that subjects may respond to compounds rather than components. Concern with this question led to the following line of theoretical analysis.

4. A Context-Sensitive Theory of Discrimination Learning. Again description of this work will be abbreviated because it is soon to be published and copies are enclosed of an abridged presentation given at a Conference on Discrimination Learning Models.

This work grew out of the lack of support for predictions of an incentive model for learning to learn and an increased concern for the role of memory phenomena in discrimination learning, particularly context effects. The theory is an attempt to account for within problem learning in a variety of paradigms as well as to begin to predict various transfer phenomena.

The basic assumptions of the model are that:

1) The amount of association information or feedback available from a cue in a given context is reduced either by changes in the cue under consideration or its context. For a given cue on a dimension the other cues in the situation form part of the context. The similarity on a given dimension of one cue to another will be represented by a parameter whose value is between 0 and 1, with 1 corresponding to identity on that dimension and 0 corresponding to total dissimilarity. The reduction due to generalization along a particular dimension will be represented in like manner.

2) The context or cue change decrements from the various dimensions are combined in a multiplicative manner to yield a single similarity measure. Thus the difference between a white triangle on the left and a black square on the right involves a difference in position (p), brightness (b), and form (f). Any information associated with the white triangle would generalize to the black square but would be reduced by these differences. If R is the information from a reward associated with the white triangle, then pbfR would be associated with the black square. The difference in information ($\Delta I$) as a result of such a reward trial would be (1-pbf)R.

3) In a discrimination task involving two alternatives, it is assumed that the greater the difference in information between the alternatives, the faster will be the learning to discriminate between them. This assumption was introduced strictly for convenience to avoid discussing details of any one particular model. Computer simulations have been based on one specific model but many other models could well satisfy the third assumption.

The context model makes correct predictions concerning the effects of adding different kinds of relevant cues to problems and performance on simultaneous, successive, conditional discriminations as well as mixtures of these problems.

Predictions concerning transfer phenomena are developed in a book chapter in preparation. The key assumption is that the

similarity of two stimulus values on a dimension takes on one of two values, depending upon whether or not that dimension is attended to on a trial. Specifically there is less generalization (the similarity is effectively less) when a dimension is attended to than when it is not.

Space does not permit a full exploration of this assumption. The model predicts that intra-dimensional shifts will be solved faster than extra-dimensional shifts but that this effect may interact with dimensions (as has commonly been found). The model is compatible with most data concerning shift effects, backward learning curves, redundant relevant cue learning, blocking, and certain generalization phenomena. It is too early to evaluate the context model but the initial impression is that it will prove to be quite promising in analyzing children's learning.

C. Dimensional Responding and Preferences

1. Measuring and training dimensional preferences (Medin, 1973b). This experiment suggests that techniques to measure dimensional preferences may actually create them. Thirty kindergarten children were presented with one standard stimulus and pairs of comparison stimuli and asked to judge which comparison stimulus was most like the standard. The stimuli could differ in size and brightness and both free (preference) and forced choice trials were given. The first trial was always a forced trial. The children showed strong dimensional preferences and frequently made errors on forced trials involving the nondominant dimension. However, these

preferences could be almost perfectly predicted on the basis of the first forced trial, with 27 of the 30 children responding primarily to the dimension involved on the initial trial.

2. Preference in a two- versus three-choice situation. The preceding experiment indicated that children were likely to respond on the basis of a single dimension. A followup experiment compared a two- and three-choice situation where stimuli could differ in size, brightness, and saturation. The two-choice situation presented the two extremes of size, brightness, and saturation, and subjects were given intermediate stimuli and asked which of the two choice boxes they belonged to (a mailbox game). The three-choice situation involved as choices, (1) a large, low saturation, low brightness choice box, (2) a small, low saturation, low brightness choice box, and (3) a small, low saturation, high brightness choice box. Sixteen kindergarten children received 64 trials on the two-choice situation and sixteen children were tested in the three-choice situation.

For the two-choice situation, all subjects responded on the basis of a single dimension with size being dominant followed by saturation and then brightness. In contrast, 9 of 16 children in the three-choice paradigm responded on the basis of more than one dimension. Further analysis showed that these multidimensional judgments were based on sequential rather than simultaneous judgments. For example, all large stimuli might be put in the large choice box (regardless of their saturation and brightness) and then the sorting of stimuli into the two

smaller boxes might be based on saturation. No child evidenced simultaneous judgments of more than a single dimension.

3. Developmental changes in dimension discriminability. A more direct assault on the relation between stimulus analysis and learning compared the dimensional processing of 20 kindergarten and 20 third grade children on four perceptual continua.

Subjects judged the similarities of sets of colors, sizes, shapes, and length-width ratios. The similarity data were analyzed with a multidimensional scaling procedure which included both a Euclidean space for the stimuli and dimensional weights or saliences for the two age groups. For simple dimensions no age differences were evident. However, on any but the simplest dimensions (e.g. size) the results imply that older children process dimensional information more efficiently than younger children. Older children had greater weights for complex dimensions and in some cases secondary dimensions emerged for them even though the stimuli were nominally unidimensional (e.g., dimension of "simplicity" appeared in comparison of number of sides). Therefore it seems that assessing development of discrimination learning requires both a theory of perception and a theory of learning. More than just changes in relative salience occur with age. The procedure used in this experiment suggests a technique for measuring changes in perceptual processing skills.

D. Summary

In brief the project period provided a number of important findings. First, an incentive model developed by the principal

investigator received little support but a context model for learning and transfer was advanced which may represent a significant step toward achieving the goals outlined. Further analysis cast doubt on what appeared to be developmental changes in mediational skills but suggested that dimensional processing skills are modified with age. This latter finding and other experiments provide evidence that attention might best be viewed as not simply a fixed quantity to be doted out to various dimensions but that individuals may differ in the total capacity for attention. Finally, an experiment yielded preliminary evidence that learning rules with exceptions may be distinctly different from exceptionless rules.

## References

Medin, D. L. The role of reinforcement in discrimination learning set in monkeys. Psychological Bulletin, 1972, 77, 305-318.

Reese, H. W. Discrimination learning set in children. In L. P. Lipsitt & C. C. Spiker (Eds.), Advances in Child Development and Behavior. Vol. 1. New York: Academic Press, 1963.

Spiker, C. C. Performance on a difficult discrimination following pretraining with distinctive stimuli. Child Development, 1959, 30, 513-521.

Tighe, T. J., Glick, J., & Cole, M. Subproblem analysis of discrimination shift learning. Psychonomic Science, 1971, 24, 159-160.

Vurpillof, E. The development of scanning strategies and their relation to visual differentiation. Journal of Experimental Psychology, 1968, 6, 632-650.

III. Publications or Articles Contemplated

A) Papers presented or published articles

Medin, D. L. Implications of subproblem analysis of reversal and nonreversal shifts for attention theory. Presented at a Symposium on Discrimination Transfer, Mathematical Psychology Meetings, San Diego, August 1972.

Medin, D. L. A context-sensitive theory of discrimination learning. Presented at a Conference on Discrimination Learning Models, Rockefeller University, April 1973.

Medin, D. L. Measuring and training dimensional preferences. Child Development, 1973, 44, 359-362.

Medin, D. L. Subproblem analysis of discrimination shift learning. Behavior Research Methods and Instrumentation, 1973, 5, 332-336.

B) In press or submitted

Medin, D. L. Memory, context, and discrimination learning theory. To appear in Bower, G. H. (Ed.), The Psychology of Learning and Motivation, Vol. 8. New York: Academic Press, 1975.

Fahrmeier, E., Rourke, D. L., & Medin, D. L. Developmental changes in dimensional discriminability. (Submitted.)

C) Articles in preparation

Medin, D. L. Multi-dimensional sorting in two versus three choice situations.

Medin, D. L. Learning rules where there are exceptions to the rules.

## IV. Abstract

Seven main experiments and three pilot studies were conducted to develop and test theories of transfer of training in children. Initial tests failed to support an incentive model for learning and transfer but a new model given strong emphasis to the role of context in learning was developed which accounted for a wide range of learning and transfer data. Although this was clearly the most salient result, other work revealed (1) learning rules when there are exceptions to rules may be much more difficult than one might have been led to imagine, (2) simply solving any problem (learning the rule) may underlie the basis of the easy-to-hard transfer effects in children, (3) tests to measure dimensional preferences may create them, (4) young children may use more than a single dimension in making choices when decisions can be considered sequentially, (5) dimensional processing skills may increase with age in ways that are more significant than simply changes in the allocation of attention, (6) some claims concerning changes in mediating skills with age were most likely premature.

Overall the project demonstrates the feasibility of studying the modifiability of learning rates or learning styles by examining component processes and further implies that generalizations concerning development of learning skills are nearly impossible to make in a model-free manner. Finally, the project period led to the development of a model which may advance interpretation of many learning and transfer paradigms. Ultimately, the present research should lead to specific suggestions for

training procedures designed to optimize learning skills.